

## Secondary Predicates in Native and Nonnative Grammars

**Enchao Shi**

California State University,  
Northridge  
18111 Nordhoff Street  
English Department, CSUN  
Northridge, U.S.A.  
[enchao.shi@csun.edu](mailto:enchao.shi@csun.edu)

### Abstract

The typical measurement by which the nature of second language grammars is evaluated is the input of native speakers. This paper reports on data from Mandarin speakers of English ( $n = 19$ ), with an average of 10;3 (year;month) length of residence in the U.S., and native American English speakers ( $n = 19$ ), and looks at how they dealt with causatives, resultatives, and depictives under four experimental conditions. It was found that native participants did not always behave reliably; they altered, swung, and oscillated just like nonnative counterparts, and there were multiple cases where their fluctuation rates were way higher than those of the latter. Such variances were brought about by the effects of construction, task, or modality. These results cast doubt on the common practice of assessing second language grammars in terms of native intuitions and call on researchers to reconsider the assumption that second language grammars that are legitimate must be native-like.

### 1 Introduction

It goes without saying that adult second language ( $L_2$ ) learning differs from child first language ( $L_1$ ) development, owing to various identifiable disparities in cognition and maturation between the two groups. From this truism follows the logical question, for generative second language acquisition (SLA) researchers, how much of the initial state or the biologically determined precursor contributes to the acquisition of a language later in life. Previously, the issue was explored by inquiries into the developmental processes until about two decades ago when researchers started to seriously consider what it is that adults can ultimately know about the

target language that are not true of their native languages. Much in line with the developmental research, results gleaned from empirical and longitudinal studies that focus on the final  $L_2$  state, suggest, or in many cases conclude, that older learners attain different grammars than native speakers. For those born and raised in the target language setting, language development is, from the onset, controlled by UG (Universal Grammar) principles and parameters. Mature learners are subjected to all kinds of undesirable elements none of which occurs to child learners (for a review, see White<sub>a</sub>, 2003), leading them to the mastery of  $L_2$  grammars full of anomalies and aberrations.

The present experiment questions whether it is enough to measure  $L_2$  grammars against native grammars (for discussion, see Mack, 1997) and shows that native speakers vary in behavior just as much as nonnative speakers, depending on the grammatical features under analysis and the experimental conditions. The typical comparative native-nonnative studies with an eye to pinning down the biological influence bear little fruit if it is true that the final state of the model subjects lacks the supposed uniformity in the knowledge of the target grammar, which, according to Chomsky (1986, 1988, 1993), is not something in dispute. What is not being investigated is the causes of the native variations (see Shi, 2014; Shi<sub>a</sub>, in progress).

A large number of the generative  $L_2$  studies conclude that there is something amiss about  $L_2$  grammars, for their bearers deviate from the natives whose use of the target language is reliable and consistent. In an influential study involving proficient English speakers of first languages of Korean, Chinese, Indonesian, and Dutch, Schachter (1990) found that these subjects, unlike the native controls, did not always recognize errors in sentences like

\*What did Susan visit the store that had t in stock?

The extent to which an  $L_2$  group succeeded was correlated with whether its native grammar instantiated the Subjacency constraint (Chomsky, 1981) as English did. Schachter takes this as support for her Incompleteness Hypothesis; namely “incompleteness will turn out to be an essential property of any adult second language grammar” (pp. 118-119).

Johnson et al. (1996) also found from 10 Chinese speakers of English, who had on average lived in the U.S. for 6.45 years, that their abilities to recognize morphosyntactic errors from auditorily presented sentences were lower (54.2% accurate) than the native speakers (98.3% accurate). This indicates, to the researchers, that  $L_2$  grammars of the nonnative-born speakers are “not native in determinacy” (p. 343).

In an experiment on English psych verbs (*interest*, *disappoint*) and container verbs (e.g., *decorate*, *cover*), Juffs (1996) retrieved both production and comprehension data that informed him that the Chinese college students, with no living experience in an English-speaking environment, lagged behind native speakers in consistency, which tended to improve as a function of the increased proficiency level. While low- and intermediate-level learners had trouble producing or processing sentences like “The broken vase disappointed John,” those at the advanced level did as well as the English speakers.

Chen (2005), in search of the association of verbs *consider*, *find* to various complement syntactic frames, uncovered a gradient preference pattern for Mandarin speakers of English: tensed clause > infinitive clause > small clause. The finding was based on a set of within-group statistical analyses conducted to the  $L_2$  group. Assuming the lack of preference for the natives, the found preference pattern from the nonnatives suggested, to Chen, that it must be  $L_2$  grammars that were faulty, since native speakers, being native, could not go wrong.

It is not that generative  $L_2$  researchers are oblivious of or blindsided by the fact that native speakers, due to internal as well as circumstantial variables, can falter or fail to comply with the grammatical rules when called upon. For example, in the above study by Schachter (1990), she acutely noted the unusual poor performance by the English-speaking participants on the Wh-movement sentences that “had been piloted on other natives and performance has been much higher” (p. 111). As an explanation, Schachter speculated that the piloted subjects were “graduate students majoring in linguistics” (fn 19).

In Johnson et al.’s (1996) error-detection experiment, if we remove the chance responses, based on their formula (p. 343), from the native group data, its accuracy rate would drop to 96.6 percent, from the reported 98.2 percent. In the study of Chen (2005), there was a case where native speakers showed more variations, based on her computations of standard deviations, on the use of *consider/find THAT*, than nonnative speakers. But spotting such variabilities from native speakers is one thing and taking it into account is another. As has been shown time and again, native variance is largely viewed as inconsequential, reflective of the accidental glitches, and therefore dismissible.

This experiment, a mixed design, aims to do the reverse of what has been typically done; that is, to demonstrate  $L_2$  grammar is not as flawed as previously thought, provided that the random and experimental errors are carefully identified. The problem, which has been long neglected, is a methodological one – the use of native speakers as the sole yardstick to determine the nature of second language grammars. The null hypothesis tested is that speakers who acquire the target language natively therefore do not vary in linguistic competence; they as natives can always be counted on being up to par when it comes to the measurement of grammatical knowledge.

Nativeness may well be correlated with birthplace, but linguistic competence is not. To tap into the components of the faculty of language (Chomsky, 1972, p. 27; 1986, pp. 16-17; 1998, p. 115), we ought to rethink the current research procedure. One alternative being explored here is to hold off the input effects as a pernicious confounding variable, so that no subject group is at unwarranted advantage. To that end, the study tested, under contrasting conditions, a set of infrequent yet robust syntactic frames: causatives, resultatives, and depictives. The idea behind the design was that by displacing subjects from their “comfort zone” into a “leading edge” (Rispoli, 2003, p. 819), we are able to take a better look into their inner grammar proper. The independent variables of interest are three: construction (3 levels: causatives, resultatives, depictives), task (4 levels: Guided Production, Combining-Clause, Grammaticality Judgment, Interpretation Task), and modality (2 levels: production, comprehension). The construction effects are examined by holding the modality and task effects neutral. To factor in the influence of task and modality, pairs of

group data from the comparable tasks or modalities are analyzed. And finally, to see if the key variables (modality  $\times$  task) interact, tasks from different modalities are compared in pairs. The null hypothesis is rejected just in case the empirical data shows that those who speak English day in and day out fail to deliver the expected outcome on tasks presented under various conditions. They show tendencies to respond to the intricate properties of constructions, tasks, or modalities as opposed to their grammatical knowledge in ways of nonnative speakers.

### 1.1 Causatives, Resultatives, and Depictives

Accounts have been put forward for causatives, resultatives, and depictives, commonly known as secondary predicates (see Shi, 2003; Hale and Keyser, 2002; Levin and Rappaport Hovav, 1995; Jackendoff, 1990; Chomsky, 1981). The underlying framework for the present experiment is a combination of Distributed Morphology (Matushansky and Marantz, 2013; Embick and Noyer, 2007; Harley and Noyer, 2003; Marantz, 1997) and Lexical Argument Structure (Hale and Keyser, 1993, 2002). (1) illustrates the three constructions under analysis:

- (1) a. The stories about animals interested Mary.  
 b. Sally could have complained herself calm.  
 c. The invited speaker delivered the speech drunk.

(1a) is a causative in contrast to a periphrastic structure, in that it consists of a single tensed predicate whereas the latter two predicates (*The stories about animals made Mary interested*). Mandarin for the most part allows the bi-clause causatives.

As a resultative, (1b) comprises a main verb and a secondary (resultant) predicate, as shown in (2):

- (2) [<sub>V</sub> V<sup>0</sup> [<sub>AP/PP</sub> DP A<sup>0</sup>/P<sup>0</sup>]]

Semantically, a resultative expresses a cause-event leading to a result-event (Rothstein, 2006; Rappaport Hovav and Levin, 2001; Simpson, 1983, 2006; Washio, 1997; Hoekstra, 1988, 1992; Napoli, 1992; Roberts, 1988; Williams, 1980; Green, 1973; Halliday, 1967). Resultatives in the study fall into

seven subtypes classified based on the syntactic categories of the main verbs, which are given in (3).<sup>1</sup>

- (3) a. The defendant kicked the victim unconscious.  
 b. The little boy ate himself sick.  
 c. George joked himself out of his job.  
 d. Sally would sleep her headache away.  
 e. The waiter could quickly wipe the water off the table.  
 f. The sodas broke open.  
 g. The hiker followed the stars out of the forest.

(3a) and (3b) both are headed by transitive verbs except that “eat” can alternate as an intransitive. Verbs in (3c, d), inherently monadic, take a fake NP (3c) or an unselected DP (3d) as part of the secondary predication. “Wipe” of (3e), a two-place predicate, takes an unconventional DP “the water.” (3f) involves an unaccusative “break,” where the surface DP is the subject of the secondary predicate “open.” Unlike the rest, (3g) is a subject resultative, in that the abstract subject of the secondary predication PRO co-refers to the main-clause subject.

Mandarin resultatives splits into *de*-resultatives and V-V resultatives (see, e.g., Huang et al., 2009; Huang, 2006; Zhang, 2001; Cheng, 1997; Sybesma, 1997; Cheng and Huang, 1995; Zou, 1994). Neither matches the resultatives in English. While the *de*-constructions roughly correspond to the English canonical, bi-clausal resultatives (*The defendant kicked the victim until he became unconscious*), the V-V structures are rarely observed in English (see Shi, in progress). Besides, Mandarin has causative resultatives (Huang, 1988) or inverted readings (Li, 1998); neither is possible for today’s English.

Depictives (1c) is distinct from resultatives; a secondary resultative functions as a complement to the matrix verb, a depictive an adjunct that is predicated of an argument of the primary predication (for differences, see Rothstein, 2006; Schultze-Berndt and Himmelmann, 2004; Rapoport, 1999; Stowell 1991; Hoekstra, 1988; Halliday, 1967). For convenience, let us take a depictive construction to be (4),

- (4) [ DP<sub>i</sub> [<sub>VP</sub> V DP<sub>j</sub>] [PRO<sub>i/j</sub> DepP ]]

<sup>1</sup> These are the actual resultatives produced under GP condition by Mandarin speakers of English, with minor modifications like use of a pronoun instead of a proper noun.

where indexing means co-reference. Two types of depictives were investigated: subject depictives (5a) and object depictives (5b), and with an unaccusative main verb, the object was fronted (5c).

- (5) a. The invited speaker<sub>i</sub> delivered the speech drunk<sub>i</sub>.  
 b. She bought the furniture<sub>i</sub> unpainted<sub>i</sub>.  
 c. The package<sub>i</sub> arrived broken<sub>i</sub>.

A depictive is well-formed if the attribute identified by the depictive phrase holds at the time of the main event, or (5c) is false if it means something other than *the package was broken when it arrived*.

Depictives as an independent secondary predicate has not found its way into Mandarin (see Zhang, 2001, for a different view), although it does occasionally show depictive elements (for distinction, see Himmelmann and Schultze-Berndt, 2006).

In terms of the abstractness determined by the frequency effect, we see a hierarchy for English: causatives > resultatives > depictives, where the least frequent and hence the least accessible is the last. Chinese differs: resultatives > causatives > depictives. Mandarin resultatives is far more productive than English resultatives (Huang et al., 2009; Huang, 2006; Li, 1998; Sybesma, 1997), its causatives is substantially less so (Thompson, 1973), and depictives is merely absent (Shi, in progress, 2003).

## 2 Method<sup>2</sup>

### 2.1 Participants

Nineteen Mandarin speakers of English (9 males, 10 females), between the ages of 26 and 48 ( $M = 37.3$ ), were tested, along with nineteen native-born American college students<sup>3</sup> (4 males, 15 females), between the ages of 17 and 41 ( $M = 27.4$ ). The nonnatives were recruited based on a set of criteria, including a consecutive period of 5 living years in the U.S. and a college education. As it turned out, they had an average of over 10 years living experience (range 5-17) and were employed in the mainstream workplace in America. All participated voluntarily.

<sup>2</sup> The data reported here is part of a more comprehensive experiment (see Shi, 2003).

<sup>3</sup> One native participant was a college graduate.

### 2.2 Tasks

Participants were subjected to two production and two comprehension tasks. For each of the 30 Guided Production (GP) test items, they first read a narrative of about 3-line long and then answered a question, using words provided, in all possible ways (e.g., *The chef boiled the lobster alive; The lobster was alive when the chef boiled it*). On the Clause-Combining (CC) task (30 items), subjects converted bi-clauses (*Sam drank until there is nothing left in the bottle of whisky*) into mono-clauses (*Sam drank the bottle empty*), using key words given. Of the two, GP was more demanding, given that subjects were asked to produce multiple answers. Also, for CC, they were allowed to leave a question blank.

The Grammaticality Judgment (GJ) task asked subjects to assess a total of 81 items for grammaticality on a scale of -3 to +3, with zero = no judgment. The analyses given below were based on 47 (2 causative, 2 inchoative, 16 resultatives, 13 canonical resultatives, 8 depictives, 6 canonical depictives) and the rest were fillers (ill-formed sentences) or sentences that turned out to be structural ambiguous, which then did not enter into the analyses. In choosing a numeral other than +3, subjects were instructed to identify the problem site by underlining the relevant word(s). The Interpretation task (IT) is the mirror image of the CC task. Participants matched mono-clauses with bi-clauses as paraphrases. Some items were 3-way ambiguous (resultative, object depictive, subject depictive) (6 items), others 2-way ambiguous (resultative and object depictive) (3 items), and still the others 1-way ambiguous (resultative) (5 items). As was designed, IT was relatively more challenging than GJ since multiple semantic recognitions forced participants to reconstruct more than one underlying representation.

The experiment tested 105 verbs or verb-pairs under three to four conditions and the analyses here are based on 96 of them: 6 causative verbs, 60 resultative verb-pairs, 30 depictive verb-pairs.<sup>4</sup> All tasks were individually administered in a paper and pencil format, with no time limit. Subjects were requested to carry out the tasks on their own.

<sup>4</sup> Three depictive verb-pairs tested under GP and four under CC were removed from the analyses for being potentially interpretable as conditionals or concessive/causals. Two causative inchoative verbs tested under GJ were also removed.



## 2.3 Results

Previously, it has been attested in study after study that while the natives fell victim to random errors, the nonnatives erred systematically (e.g., Kweon and Bley-Vroman, 2011; Chen, 2005; Papp, 2000; Johnson et al., 1996; Juffs, 1996). Of all the plausible explications, one that stands out in particular, albeit rarely noted, is that such native-nonnative disparities reflect more of the properties of the linguistic variables being tested, which inadvertently give the monolinguals an unfair head start. So it would be not just interesting but essential to see whether the presumed perfect or near-perfect native performance still prevails in the absence of such advantages.

Causatives, resultatives, and depictives are posited to be the linguistic features, which are vibrant in positive evidence and yet low in frequency, hence providing us with a unique testing ground for the native systematic variance and the native-nonnative difference.

### 2.3.1 Causatives

The causative constructions were studied under three conditions. Table 1 enumerates the means and sigmas (standard deviations *s.d.*), for the groups, of the verbs tested: *disappoint*, *interest* (GP), *bore*, *frustrate* (CC), *lengthen*, *awake* (GJ). Based on a

	<u>GP</u>		<u>CC</u>		<u>GJ</u>	
	<i>M</i>	$\sigma$	<i>M</i>	$\sigma$	<i>M</i>	$\sigma$
NS	1.79	.42	1.42	.77	11.26	1.94
NNS	1.63	.6	.95	.62	10.11	2.11

Table 1. Causative Means and *s.d.*

one-factor ANOVA with repeated measures, conditional variabilities were robustly found from both groups. The native speakers (NS) shifted in performance, beyond chance, from task to task,  $F(2, 56) = 376.4$ ,  $p < .05$ , partial  $\eta^2 = .96$  and so did the nonnative speakers (NNS),  $F(2, 56) = 267.15$ ,  $p <$

.05, partial  $\eta^2 = .94$ . Interestingly, between the two, it was the NS that wobbled at a higher rate. Regardless, for both groups over 94% of the total behavioral variance was caused by the general task effects. This evidence strongly suggests that knowing causatives does not always guarantee its use, which is true of every subject irrespective of where he or she was born. Take the native speakers as an example. They all (100%) composed the target causative sentence using *disappoint* under GP, but only 58% did so using *frustrate* under CC.

No significant difference was found between the groups for any given task, based on three independent-sample *t*-tests,<sup>5</sup> with the alpha being set at .02 or one-third usual .05 alpha to offset alpha inflation. This evidence doubtless is unfavorable to Juffs's findings, according to which Chinese participants should have flunked no matter what conditions they were tested under.

### 2.3.2 Resultatives

In Table 2, group averages are given of the resultative data elicited from four tasks. If a group behaves

	<u>GP</u>	<u>CC</u>	<u>GJ</u>	<u>IT</u>
	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>
NS	7.11 (1.76)	13.37 (1.21)	78.21 (9.07)	10.84 (1.46)
NNS	3.37 (1.7)	12.05 (2.25)	76.68 (10.78)	11.89 (1.37)

Table 2. Resultative Means and *s.d.* (in brackets)

as though it is controlled by the experimental conditions under which it is measured, then a large *F* value arises from a one-way repeated measures ANOVA. This is precisely what was in fact obtained,  $F(3, 54) = 1020.41$ ,  $p < .05$ , partial  $\eta^2 = .98$  ( $L_1$  group), and  $F(3, 54) = 713.58$ ,  $p < .05$ , partial  $\eta^2 = .98$  ( $L_2$  group). The findings clearly showed that neither group was good at breaking the conditional barriers; all subjects responded in accordance with

<sup>5</sup> All *t*-tests conducted in the experiment were two-tailed.

the task intricacies as opposed to what they knew about the target language. To illustrate, the native participants created 5 resultatives using *eat-sick* (*The boy ate himself sick*) under GP, compared with 14 using *work-to death* (*Mark worked himself to death*) under CC, despite the fact that both main verbs came from the same categorial class a.k.a transitive and selected the same type of secondary predicates. Had the native grammars been as steady and fast, unaffected by the circumstantial vagaries, as has been shown again and again in the literature, we would not have seen variations in performance of such magnitude.

Equally surprising is the finding that of the four conditions, the monolinguals outperformed the bilinguals only under GP,  $t(36) = 3.35$ ,  $p < .01$  (near one-fourth of normal .05 alpha), and this native merit was cancelled out under IT, where the reverse was found,  $t(36) = 2.39$ ,  $p > .01$ .<sup>6</sup> This rather unexpected result could be explained away as an effect of positive transfer, but this possibility diminishes in face of the aforementioned differences in Mandarin-English resultatives. Additionally, it would leave unexplained the native-nonnative congruence attested under the CC and GJ conditions ( $p > .01$ ).

### 2.3.3 Depictives

If the notion that a legitimate  $L_2$  grammar must be native-like is sound, then depictives gives us reasons to contemplate the possibility that it is not. Due to the effects of input frequency and crosslinguistic differences, persistent  $L_2$  aberrations should readily come along, parting from the native benchmark. This prediction has not quite panned out, as seen in

	<u>GP</u> <u>M</u>	<u>CC</u> <u>M</u>	<u>GJ</u> <u>M</u>	<u>IT</u> <u>M</u>
NS	4 (1.41)	2.74 (1.28)	40.32 (4.32)	6.32 (2.38)
NNS	2.58 (1.77)	2.53 (1.26)	37 (6.86)	3.11 (1.29)

Table 3. Depictive Means and *s.d.* (in brackets)

Table 3. The depictive data, in contrast to the resultative data, appeared to be more homogeneous, with smaller standard deviations across groups. But a one-factor ANOVA for repeated measures ascertained that this visual impression was not what it seemed. Both groups demonstrated sensitivities to the challenges imposed by individual tasks, causing them to behave chancily. As before, it was the natives that were plagued by such unwarranted variabilities,  $F(3, 54) = 860.56$ ,  $p < .05$ , partial  $\eta^2 = .98$  (natives);  $F(3, 54) = 435.96$ ,  $p < .05$ , partial  $\eta^2 = .96$  (nonnatives). A series of two-sample *t*-test revealed that the two groups differed drastically under GP ( $p = .0097$ ) and IT ( $p = .000$ ), but they were indistinguishable under CC ( $p = .61$ ) and GJ ( $p = .08$ ).

As shown in Table 1, 2 and 3, the native means for causative, resultative and depictive productions under CC are 1.42, 13.37, and 2.74 and the nonnative means are .95, 12.05, and 2.53. According to a one-way repeated measures ANOVA, the natives did not treat the three constructions evenly,  $F(2, 36) = 851.09$ ,  $p < .05$ , partial  $\eta^2 = .98$ , and neither did the nonnatives,  $F(2, 36) = 306.71$ ,  $p < .05$ , partial  $\eta^2 = .95$ . The results indicate that not only did the conditions contribute to variable behaviors but also the constructions. With everything else being equal, one thing that is quite clear is that in assessing whether nonnative grammars are up to native par, one should take as little risk as possible of overlooking the effects of tasks and linguistic variables. Studies that hinge on a single trial or banal linguistic features undermine both their internal and external validity (see Cook and Campbell, 1979). They most likely fail to shed light on the research questions under probe, let alone be fit to generalize beyond the data at hand.

### 2.3.4 Variations Within or Between a Modality

What has been presented is the overall group variances in the production and comprehension of causatives, resultatives, and depictives. What happens if the modality effect or the task effect is partialled out? At the minimum, it is of import to know whether performance disparities would remain when tasks are isolated from the modality so that they do not covary, and if the two variables interact, how it proceeds across the two groups.

<sup>6</sup> The between group difference was a bit short of the critical value of 2.43.

Let us first look at causatives. This time we focus on the production data and see if there is still task-induced variation without the comprehension mode as a covariate. Four causative verbs were analyzed: *disappoint*, *interest*, *bore*, *frustrate*, under GP and CC. The NS group achieved means of 1.79 (GP) and 1.42 (CC) and the NNS group means of 1.63 (GP) and 0.98 (CC) (Table 1). The production tasks were found not to affect the groups in the same way, based on two paired *t*-tests. The natives, whether to create causatives from scratch (e.g., *The stories about animals interested her*) or from bi-clauses (e.g., *The lecture was so long that Jack became bored* → *The lecture bored Jack*), performed consistently,  $t(18) = 1.9$ ,  $p > .025$ , but the nonnatives did not,  $t(18) = 3.34$ ,  $p < .025$ . Note that the two groups did not deviate or coincide across tasks to the same extent; the Mandarin speakers fell way below expectations under CC. Among all other plausible culprits, one that seems particularly relevant is that this was a lexical problem, as identified by Pinker (1989); namely, lexical knowledge tends to vary greatly from speaker to speaker, “no two alike” (p. 2). Under CC, for example,  $L_2$  subjects succeeded 14 times (74%) with *bore*, compared to 4 times (21%) with *frustrate*.

In the case of resultatives, a broad modality-within, task-related difference was found. The natives as well as the nonnatives excelled under the CC condition, compared with the GP condition according to a set of paired *t*-tests ( $p < .025$ ). All subjects, regardless of whether English was their first or second language, had a higher success chance to combine (6a) into (6b) than to ab initio construct (6c, d) based on brief narratives. The same was

- (6) a. Margaret screamed and as a result she became sore in her throat.  
 b. Margaret screamed herself hoarse.  
 c. Sara read herself angry.<sup>7</sup>  
 d. Bill could have watered the plants flat.<sup>8</sup>

found for resultatives under the comprehension mode; subjects, native and nonnative, did substantially better under the less stressful GJ condition, than the more stressful IT condition ( $p < .025$ ).

<sup>7</sup> For (6c), no sample was obtained from either of the groups.

<sup>8</sup> Three from  $L_1$  group yielded (6d), out of a total of 19, whereas none from  $L_2$  group succeeded.

The effects of tasks relative to depictives were a bit murky for both of the groups. For the natives, though they used distinct strategies in dealing with tasks of a given modality at  $p = .025$ , the putative influence was not forthcoming. In comprehension, for example, they judged, as anticipated, better on GJ than IT, but in production they were more successful ( $p < .025$ ) at constructing depictives without cues (7a, b) than modifying a canonical depictive (7c) into a depictive (7d). The experimental group

- (7) a. Robert opened the window wet.<sup>9</sup>  
 b. The package arrived broken.<sup>10</sup>  
 c. Bob sold his car when it was new.  
 d. Bob sold his car new.

showed no production-related task effect ( $p > .025$ ). For the comprehension tests, they met the expectation, being more accurate on GJ than IT ( $p < .025$ ). All statistics were based on paired *t*-tests with alpha set at .025.

We have by far seen 16 cases where tasks alone, implemented in identical or different modes, either enhanced or inhibited the activation of the sought grammatical knowledge. The other 2 cases showed no task effect in production, one concerning the NS group that treated causatives under GP and CC blindly and the other the NNS group that handled depictives under GP and CC indiscriminately. Out of the total 16 cases of task-related variations, the natives showed more variances in 5 cases, whereas the nonnatives showed only in one case.

This opens up a crucial question of whether a task effect still holds across modalities or whether task and modality interact. To see this, we reexamined the resultative data collected in the CC production mode and the IT comprehension mode. Two paired-sample *t*-tests were conducted. Results show that the English-speaking subjects were more sensitive to the effects of task and modality,  $t(18) = 5.9$ ,  $p < .025$ ; they were more proficient at producing than identifying resultatives. By contrast, the Mandarin speakers of English were indifferent, whether to combine bi-clauses into resultatives or to match target sentences with resultative readings,  $t(18) = .33$ ,  $p > .025$ .

<sup>9</sup> For (7c), only two subjects from each group provided the target depictive construction.

<sup>10</sup> For (7d),  $L_1$  group outstripped  $L_2$  group; it collectively created 9 depictives in contrast to 3 by the nonnative group.

This result is at odds with the widely reported findings on two fronts. First, if variable behavior was ever found as a function of tasks, it should be retrieved from the second language learners only, as has been shown in Chen (2005), Kong (2005), White (2003b), Lardiere (1998a, 1998b), Johnson et al. (1996), Sorace (1993). Natives by definition are sticklers for grammatical rules. Second, if a group could care less about whether to produce or to judge, it must be the one whose members speak English their whole lives. What was uncovered in the present study is the opposite; it was the nonnatives that were more of rule enforcers than the natives in composing or parsing resultatives.

The finding that the native-born subjects were likely to fluctuate could be confirmed provided that the similar depictive data was attested. To that end, let us compare the relevant data under CC and IT through two paired *t*-tests. As it turned out, only the natives performed in an unbalanced manner; they, while performing at a rate of 70% in recognizing depictive readings, did so at a rate of 46% in recreating depictives,  $t(18) = 5.9, p < .025$ . In contrast, the nonnatives barely altered between the two modalities,  $t(18) = .33, p > .025$ . To reiterate, it was not once but twice that the native controls showed mixed performances – constructing more than identifying resultatives, but identifying more than constructing depictives. The same failed to be found from the nonnatives. This suggests the possibility that the receptive vs. productive knowledge was not exactly the same insofar as the two groups were concerned.

To summarize, both the experimental and the control groups were found to vary along the lines of construction, task, and modality. The overall task effects were found, with mixed between-group differences. For causatives, all subjects shifted in performance across tasks and the natives did so to a greater extent. Within a given task, no difference whatsoever was ever found between the groups, under both modalities. On resultatives, both groups were identified with the similar task-based variations. Between the two, the natives prevailed under GP and the nonnatives under IT, but the two did not differ under CC and GJ. For depictives, similar task-related variance was found from both  $L_1$  and  $L_2$  groups, but the former outstripped the latter under the GP and IT conditions.

When holding the modality constant, subjects still varied across tasks. On resultatives, both groups

behaved variably between tasks under a single modality. This pattern was nevertheless not found for depictives. Only the native participants were observed to switch strategies between GP and CC. On the comprehension side, the controls did better under GJ than IT and, the reverse was true of the experimental group. For the interactive effects, we see that the English speakers produced under CC more resultative samples than identified under IT, whereas the Mandarin speakers treated them all the same. On depictives, the same asymmetric performance pattern was found for both groups. Like the natives who did better under IT than under CC, the nonnatives showed the same pattern.

### 3 Conclusion

The major finding of the study is that speakers of English, as an  $L_1$  or  $L_2$ , are not unsusceptible to the variability problem. By measuring the linguistic knowledge and the extraneous factors under which such knowledge is elicited, it shows the ties of the failures of activating knowledge to the unduly interference of tasks, constructions, or modalities. Contra the previous research, native speakers, just like their nonnative counterparts, are found to shift linguistically, not occasionally but most of the time. Where the bilinguals are spared, the monolinguals still succumb to the modality effects in both the resultative and depictive cases. Given all this, it is hard not to reject the null hypothesis and argue that native speakers do vary in linguistic competence. Generative  $L_2$  researchers have barely paid attention to this aspect, whose only interest seems to be in the extent to which  $L_2$  grammars correspond to  $L_1$  grammars. The native-like requirement is difficult to reconcile with the following two facts: (1) the natives alter, swing, oscillate to a greater extent than the nonnatives; (2) nonnative grammars different from native grammars are still permitted in UG. This should be enough for us to rethink about the widely accepted research practice whereby  $L_2$  grammars are assessed exclusively through the lens of the native norm or what Mack (2003) calls the *monolingual-comparison approach*. Instead, second language grammars should be, first and foremost, evaluated with respect to UG principles and operations and the input effects.  $L_2$  intuitions, no matter how nonnative-like, could still tell us about the UG involvement in adult  $L_2$  development.



## References

- Marina Yueh-cheng Chen. 2005. English Prototyped Small Clauses in the Interlanguage of Chinese/Taiwanese Adult Learners. *Second Language Research* 21(1): 1-33.
- Lisa L-S Cheng. 1997. Resultative Compounds and Lexical Relational Structure. *Chinese Languages and Linguistics III: Morphology and Lexicon*. Taiwan.
- Lisa L-S Cheng and C.-T. James Huang. 1995. On the Argument Structure of Resultative Compounds. In M. Chen and O. Tzeng (Eds.), *In Honor of William Wang: Interdisciplinary Studies on Language and Language Change*, 187-221. Pyramid Press, Taipei.
- Noam Chomsky. 1972. *Language and Mind* (Enlarged Edition). Harcourt Brace Jovanovich, San Diego.
- Noam Chomsky. 1981. *Lectures on Government and Binding*. Foris, Dordrecht.
- Noam Chomsky. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. Fraeger Publishers, Westport, CT.
- Noam Chomsky 1988. *Language and Problems of Knowledge: The Managua Lectures*. MIT Press, Cambridge, MA.
- Noam Chomsky 1993. *Language and Thought*. Moyer Bell, Wakefield, RI.
- Noam Chomsky. 1998. Some Observations on Economy in Generative Grammar. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky (Eds.), *Is the Best Good Enough?* 115-128. MIT Press, Cambridge, MA.
- Thomas D. Cook and Donald T. Campbell. 1979. *Quasi-experimentation: Design Analysis Issues for Field Settings*. Houghton Mifflin, Boston.
- David Embick and Rolf Noyer. 2007. Distributed Morphology and the Syntax-Morphology Interface. In Gillian Ramchand and Charles Reiss (Eds.), *The Oxford handbook of linguistic interfaces*, 289-324. Oxford University Press, Oxford, UK.
- Georgia M. Green. 1973. A Syntactic Syncretism in English and French. In B. B. Kachru, R. B. Lees, Y. Malkiel, A. Pietrangeli, and S. Saporta (Eds.), *Issues in Linguistics: Papers in Honor of Henry and Renee Kahane*. University of Illinois Press, Chicago.
- Ken Hale and Jay Keyser. 1993. On Argument Structure and the Lexical Expression of Syntactic Relations. In Ken Hale and Jay S. Keyser (Eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 53-109. MIT Press, Cambridge, MA.
- Ken Hale and Jay Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. MIT Press, Cambridge, MA.
- Heidi Harley and Rolf Noyer. 2003. Distributed Morphology. In Lisa Cheng and Rint Sybesma (Eds.), *The Second Glot International State-of-the-Article Book: The Latest in Linguistics*, 467-496. Mouton de Gruyter, Berlin.
- M. A. K. Halliday. 1967. Notes on Transitivity and Theme in English, Part I. *Journal of Linguistics*, 3(1): 37-81.
- Nikolaus P. Himmelmann and Eva F. Schultze-Berndt. 2006. Issues in the Syntax and Semantics of Participant-oriented Adjuncts. In Nikolaus P. Himmelmann and Eva F. Schultze-Berndt (Eds.), *Secondary Predication and Adverbial Modification: The Typology of Depictives*, 1-68. Oxford University Press, Oxford, UK.
- Teun Hoekstra. 1988. Small Clause Results. *Lingua*, 74 (203): 101-39.
- Teun Hoekstra. 1992. Aspect and Theta-Theory. In I. M. Roca (Ed.), *Thematic Structure: Its Role in Grammar*, 145-74. Mouton de Gruyter, Berlin.
- Ray Jackendoff. 1990. *Semantic Structures*. MIT Press, Cambridge, MA.
- C.-T James Huang. 1988. Wǒ Pǎo de Kuài and Chinese Phrase Structure. *Language*, 64(2): 274-311.
- C.-T James Huang. 2006. Resultatives and Unaccusatives: A Parametric View. *Bulletin of the Chinese Linguistic Society of Japan* 253: 1-43.
- C.-T James Huang, Y.-H Audrey Li, and Yafei Li. (2009). *The Syntax of Chinese*. Cambridge University Press, Cambridge, UK.
- Jacqueline S. Johnson, Kenneth D. Shenkman, Elissa L. Newport, and Douglas L. Medin. 1996. Indeterminacy in the Grammar of Adult Language Learners. *Journal of Memory and Language*, 35(3): 335-352.
- Alan Juffs. 1996. Semantics-Syntax Correspondences in Second Language Acquisition. *Second Language Research*, 13(2): 177-221.
- Stano Kong. 2005. The Partial Access of Universal Grammar in Second Language Acquisition: An Investigation of the Acquisition of English Subjects by L1 Chinese Speakers. *Journal of East Asian Linguistics*, 14(3): 227-265.
- Soo-Ok Kweon and Robert Bley-Vroman. 2011. Acquisition of the constraints on *wanna* contraction by advanced second language learners: Universal grammar and imperfect knowledge. *Second Language Research* 27(2): 207-228.
- Donna Lardiere. 1998<sub>a</sub>. Case and Tense in the 'Fossilized' Steady State. *Second Language Research*, 14(1): 1-26.
- Donna Lardiere. 1998<sub>b</sub>. Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research*, 12(4): 359-375.
- Beth Levin and Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface*. MIT Press, Cambridge, MA.
- Yafei Li. 1998. Chinese Resultative Constructions and the Uniformity of Theta Assignment Hypothesis. In Jerome L. Packard (Ed.), *New Approaches to Chinese*

- Word Formation: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese, 285-310. Mouton de Gruyter, Berlin.
- Molly Mack. 1997. The Monolingual Native Speaker: Not a Norm, but Still a Necessity. *Studies in the Linguistics Sciences*, 27(2): 113-146.
- Molly Mack. 2003. The Phonetic Systems of Bilinguals. In Marie T. Banich and Molly Mack (Eds.), *Mind, Brain, and Language*, 309-349. Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey.
- Alec Marantz. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. In Alexis Dimitriadis (Ed.), *Proceedings of the 27<sup>th</sup> Annual Penn Linguistics Colloquium*. Pennsylvania Working Papers in Linguistics 4, Philadelphia.
- Ora Matushansky and Alec Marantz. 2013. *Distributed Morphology Today*. MIT Press, Cambridge, MA.
- Donna Jo Napoli. 1992. Secondary Resultative Predicates in Italian. *Journal of Linguistics*, 28(1), 53-90.
- Szilvia Papp. 2000. Stable and Developmental Optionality in Native and Non-native Hungarian Grammars. *Second Language Research*, 16(2): 173-200.
- Steven Pinker. 1989. *Learnability and Cognition*. MIT Press, Cambridge, MA.
- Tova R. Rapoport. 1999. Structure, Aspect, and the Predicate. *Language*, 75(4): 653-77.
- Malka Rappaport Hovav and Beth Levin. 2001. An Event Structure Account of English Resultatives. *Language*, 77(4): 766-797.
- Matthew Rispoli. 2003. Changes in the Nature of Sentence Production During the Period of Grammatical Development. *Journal of Speech, Language, and Hearing Research*, 46(4): 818-830.
- Ian Roberts. 1988. Predicative APs. *Linguistic Inquiry*, 19(4): 703-10.
- Susan Rothstein. 2006. Secondary Predication. In Martin Everaert and Henk van Riemsdijk (Eds.), *The Blackwell Companion to Syntax, Volume I*, 209-233. Blackwell, UK.
- Jacquelyn Schachter. 1990. On the Issue of Completeness in Second Language Acquisition. *Second Language Research*, 6(2): 93-124.
- Eva Schultze-Berndt and Nikolaus P. Himmelmann. 2004. Depictive Secondary Predicates in Crosslinguistic Perspective. *Linguistic Typology* 8(1): 59-131.
- Enchao Shi. 2003. *Second Language Grammar and Secondary Predication*. Tucson: University of Arizona Ph.D. Dissertation.
- Enchao Shi. 2014. Optionality in Second Language Grammar. In C.-T. James Huang and Feng-his Liu (Eds.), *Peaches and Plums*, 353-378. Academia Sinica, Taipei.
- Enchao Shi<sup>a</sup>. (in progress). *Nativeness and Second Language MiniG*. California State University, Northridge.
- Enchao Shi<sup>b</sup>. (in progress). *Subject Resultatives and Mandarin Clause Structure*. California State University, Northridge.
- Jane Simpson. 1983. Resultatives. In Lori Levin, Malk Rappaport, and Annie Zaenen (Eds.), *Papers in Lexical-Functional Grammar*, 143-57. Indiana, Bloomington: Indiana University Linguistics Club.
- Jane Simpson. 2006. Depictives in English and Warlpiri. In Nikolaus P. Himmelmann and Eva F. Schultze-Berndt (Eds.), *Secondary Predication and Adverbial Modification: The Typology of Depictives*, 69-106. Oxford University Press, Oxford, UK.
- Antollena Sorace. 1993. Incomplete vs. Divergent Representations of unaccusativity in Near-Native Grammars of Italian. *Second Language Research*, 9(1): 22-48.
- Tim Stowell. 1991. Small Clause Restructuring. In Robert Freidin (Ed.), *Principles and Parameters in Comparative Grammar*, 182-218. MIT Press, Cambridge, MA.
- Rint Sybesma. 1997. Why Chinese Verb-Le Is a Resultative Predicate. *Journal of East Asian Linguistics*, 6(3): 215-261.
- Sandra A. Thompson. 1973. Transitivity and Some Problems with the Ba Construction in Mandarin Chinese. *Journal of Chinese Linguistics*, 1(2): 208-221.
- Ryuichi Washio. 1997. Resultatives, Compositionality and Language Variation. *Journal of East Asian Linguistics* 6(1): 1-49.
- Lydia White. 2003<sup>a</sup>. *Second Language Acquisition and Universal Grammar*. Cambridge University Press, Cambridge, UK.
- Lydia White. 2003<sup>b</sup>. Fossilization in Steady State L2 Grammars: Persistent Problems with Inflectional Morphology. *Bilingualism: Language and Cognition*, 6(2): 129-141.
- Edwin Williams. 1980. Predication. *Linguistic Inquiry* 11(1): 203-38.
- Niina Zhang. 2001. The Structures of Depictives and Resultative Construction in Chinese. In G. Jäger, A. Strigin, C. Wilder and N. Zhang (Eds.), *ZAS working papers in linguistics* 22: 191-221.
- Ke Zou. 1994. Resultative V-V Compounds in Chinese. *MIT working papers in linguistics* 22: 1-23. Cambridge, MA.